MATERIAL PROPERTY ESTIMATION USING NONLINEAR INVERSION

William R. Petrick and Lee H. Pearson

Thiokol Corporation
Applied Research Department M/S 244
P.O. Box 707, Brigham City, Utah 84302-0707

INTRODUCTION

Nonlinear inversion is the process of estimating material property values from a set of measured data. The map between the parameters and the measured data is a mathematical model. The specific mathematical model to be used in this study is the acoustic wave equation. This equation is linear in the field variable; however, it is nonlinear in the model parameter.

Nonlinear inversion presents two fundamental problems. First is the parameter estimation problem. This is the determination of a suite of model parameters that, when substituted into the appropriate mathematical model, will generate a theoretical solution which will fit a measured data set. The quality of the data fit is measured according to some preselected criterion, referred to as the objective function. The objective function could specify minimizing the squared error or, perhaps, the absolute value of the error. The second problem, which is termed the uniqueness problem, involves determining confidence intervals on the calculated set of model parameters. Of these two problems, uniqueness is probably the most important and difficult.

We have implemented a particularly stable form of nonlinear inversion known as Ridge Regression or Marquardt's method. The technique allows us to obtain inverse solutions which are more stable than standard steepest descent or Newton-Gauss type algorithms. Although this implementation of Marquardt's method guarantees stability, it does not guarantee uniqueness. The solution obtained is, among other things, a function of the initial guess model parameters. It is possible that the minimum to which the algorithm converges is not the global minimum. It may be only one of many possible local minima. There is no general way, of which we are aware, to determine the global minimum for nonlinear problems.

PARAMETER ESTIMATION

The theoretical development which constitutes the foundation of linear and nonlinear inversion may be found in texts such as Bard [1] or Menke [2]. The mechanics of parameter estimation are as follows.

Review of Progress in Quantitative Nondestructive Evaluation, Vol. 9
Edited by D.O. Thompson and D.E. Chimenti
Plenum Press, New York, 1990

Given:

G_{calc} = The calculated response from the forward model,

G_{meas} = The measured data set,

$\Delta G = G_{meas} - G_{calc}$, an n length vector, and

A = The Jacobian matrix, an n x m matrix of derivatives of the forward model with respect to the model parameters.

Calculate:

ΔP = An m length parameter change vector.

This is done by forming the matrix relationship,

$$\Delta G = A\Delta P + \varepsilon, \tag{1}$$

and solving for ΔP as follows,

$$\Delta P = (A^T A + \lambda I)^{-1} A^T \Delta G, \tag{2}$$

$$P^{(i+1)} = P^{(i)} + \Delta P, \tag{3}$$

where,

ε = The error due to neglecting higher order terms,

A^T = A transpose,

λ = The Marquardt parameter,

I = The identity matrix, and

$P^{(i+1)}$ = The (i+1) iteration model parameter estimate.

The Marquardt parameter is evaluated and "optimized" for each trial parameter change vector. Mathematically, the Marquardt parameter has the effect of weighting out small eigenvalues of the A^TA system, thus stabilizing its inverse. Small values of λ cause the system to approach Newton-Gauss behavior, while large values force steepest descent behavior. To further enhance convergence, another variable, the step size, is also employed. The algorithm is given by Marquardt [3].

Eqs. (1) through (3) guarantee that each successive iteration, i.e., each value of $P^{(i+1)}$ reduces (or at least does not increase) the squared error between the observed and calculated data. This constitutes our definition of stability. In our application the objective function to be minimized is,

$$\chi^2 = \Delta G^T \Delta G/\nu, \tag{4}$$

where $\nu = n-m$ is the number of degrees of freedom. χ^2 is termed the reduced chi squared variable and is an estimate of the data variance. Fig. 1 shows a hypothetical χ^2 surface as a function of two material property values. Such a surface would be generated by measuring a data set, then systematically evaluating χ^2 over the ranges of parameters 1 and 2, shown as axes in Fig 1. Suppose we started with this problem and wanted to find values for parameters 1 and 2 which would minimize χ^2. To illustrate the solution is initial guess dependent, consider choosing as your initial guess those parameters corresponding to point IG_1 in Fig 1.

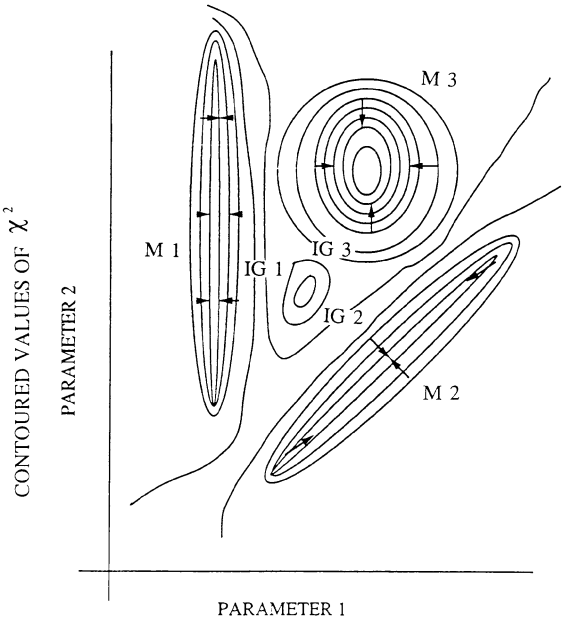

Fig. 1. Hypothetical two parameter solution space. A contour plot of squared error vs parameter values.

Since Eqs. (1) through (3) are guaranteed to find updated parameter estimates which reduce the error from the present parameter values, the algorithm would find final parameters associated with minimum M_1. Starting at point IG_2 would result in parameters at M_2, etc. When the minimum is reached which is "closest" to the initial guess there is no information regarding the existence of other minima. Marquardt's method spirals around contours, guaranteeing a decrease in χ^2 each iteration.

One might consider conducting a search over a range of parameter values, thus mapping out the positions of the minima. This approach is possible for problems with few parameters; however, if we had a data set about which we wanted to determine thirteen parameters and we only wanted to range over ten values for each parameter, we would have 10^{13} forward problems to evaluate. If it took only one second to calculate the forward problem and χ^2 for each parameter combination, we would be working on the problem for over 300,000 years. The grid search and "hedge hog" methods utilize this, somewhat trial and error, approach. A related technique is the Monte Carlo method, which differs in that the trial parameter values are obtained as the result of a statistical process, usually a random number generator. In addition to being machine intensive, these search techniques generally provide no guarantee that all minima have been found, and no parameter confidence intervals.

PARAMETER CONFIDENCE INTERVALS

Once a set of material property values $P^{(final)}$ is obtained, it is desirable to know what range of parameter values would also fit the data. Do this by defining the following manipulations, based on the Jacobian and χ^2 calculated at the final parameter values. First define a parameter covariance matrix given by,

$$cov(P) = \chi^2(A^TA)^{-1} \tag{5}$$

and a parameter correlation matrix whose elements are given by,

$$\text{cor } P_{ij} = \text{cov } P_{ij}/(\text{cov } P_{ii} \text{ cov } P_{jj})^{1/2}. \quad (6)$$

The physical meaning of each of these statistics can be understood by referring to Fig. 1 and the following explanation.

Suppose our initial guess was at point IG_1, the best fit parameters would then be somewhere in the trough at M_1. Note that in this minimum parameter 1 can have only a very narrow range of values, while parameter 2 can have a wide range of values. In this instance, the variance (a diagonal element of cov(P)) associated with parameter 1 would be small, while that associated with parameter 2 would be large.

Now suppose our initial guess was at IG_2, the best fit parameters for this initial guess would be somewhere in the trough at M_2. Note that if the value for parameter 1 increases, the value for parameter 2 must also increase in order to maintain a solution in the minimum. This parameter behavior is divulged by cor P_{12} = +1 in Eq. (6). This denotes a +1 correlation coefficient between parameters 1 and 2. A correlation coefficient near +1 indicates that only the ratio between two parameters can be determined. A correlation coefficient near -1 indicates that only the product of two parameters can be determined. To describe a -1 correlation coefficient the minimum at M_2 would be rotated 90 degrees.

If our initial guess was at point IG_3, the best fit parameters would be found at M_3. This is the best of all situations; both parameters would have small standard deviations and the correlation coefficient would be near zero.

It should be restated that once a minimum is found there is no information regarding the existence of any other minimum. The statistics described above refer only to the parameter combinations at the minimum being occupied. For a nonlinear problem, there is generally no way to determine how many minima exist. Unlike the examples shown in Fig. 1, the minima may also be highly contorted. Linear problems will have only one minima.

ILLUSTRATIVE EXAMPLE: INVERSION OF A PULSE ECHO TIME SEQUENCE

It is our aim to illustrate the importance of parameter confidence intervals. The need for this measure of uniqueness will be apparent when we show that a wide range of material property values are capable of fitting a given data set.

A set of pulse echo ultrasound data was collected from a test specimen composed of a Lucite plate clamped to a nylon plate. Between the plates was a thin layer of water. An appropriate forward solution for this combination of source and material geometry is a normally incident, plane, compressional wave in layered media. This solution has two important components which contain material property information. The first is the reflection coefficient which exists at each material property interface (Eq. 7). The second is a phase factor, or time delay term which accounts for the travel time in each layer (Eq. 8).

$$(\rho_2\alpha_2 - \rho_1\alpha_1)/(\rho_2\alpha_2 + \rho_1\alpha_1) \quad \text{Reflection coefficient.} \quad (7)$$

$$e^{-i(1/\alpha - ic)t\omega} \quad \text{Phase shift.} \quad (8)$$

The material properties which interest us are velocity (α_i), density (ρ_i), thickness (t_i), and attenuation (C_i) in each layer.[1] The ability to apply constraints is important in any successful inversion algorithm. Constraints allow the practitioner to input known information about a set of materials. Examples of constraints are; not allowing certain parameters to become negative, forcing the total sum of certain parameters to be a particular value, holding parameters constant, etc.

Inversion Results

Lab measurements were made of the velocity, thickness, and density for the Lucite and nylon plates. We did not measure attenuations for these materials. A thin layer of water filled the contact region between the two plates. Measured values were:

Lucite	α = 0.27 cm/μsec	ρ = 1.22 gm/cc	t = 0.58 cm
nylon	α = 0.27	ρ = 1.15	t = 1.45

Four different sets of parameters were used for initial guesses. These encompassed both two and three layer models. The three layer models explicitly contained a layer that could account for the water filled contact. Fig. 2 shows the convergence for the various models. As depicted in Fig. 2, virtually all of these models converged to roughly the same RMS data fit error. Notice uniform convergence for all of the models. A summary of the different models is as follows:

Model A: A three layer model in which densities and thicknesses were held constant at their measured values for all layers except the thin water layer. The initial guess material properties were the measured values. The raw data and the data fit are shown in Fig. 3. The final model was:

Lucite	α = 0.27	ρ = 1.22*	t = 0.58*	C = 0.005
contact	α = 0.08	ρ = 1.00*	t = 0.001	C = 0.000
nylon	α = 0.27	ρ = 1.15*	t = 1.45*	C = 0.01

The asterisk indicates the parameter was held constant.

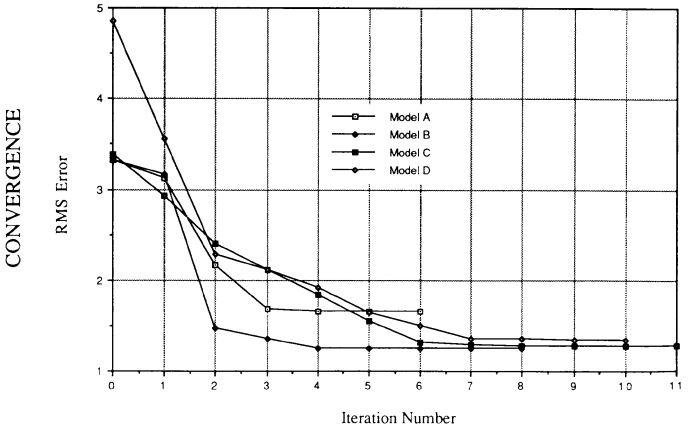

Fig 2. Convergence for four different models.

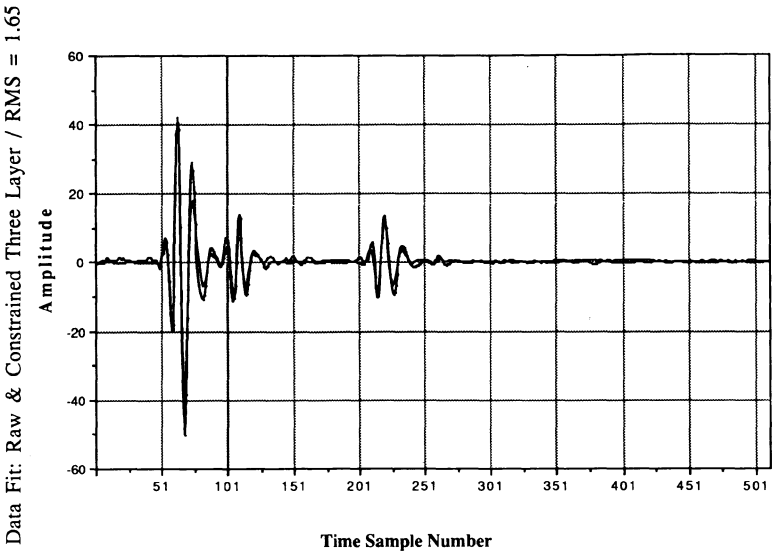

Fig. 3. Model A data fit.

Model B: A three layer model with the same material property values for initial guesses as used in Model A but this time allowing all material property values to change. The raw data and the data fit are shown in Fig. 4. The data fit is slightly better than the Model A fit. This is to be expected since Model B has a greater number of degrees of freedom; i.e., more parameters to adjust. The final model was:

Lucite	$\alpha = 0.28$	$\rho = 1.04$	$t = 0.57$	$C = 0.001$
Contact	$\alpha = 0.09$	$\rho = 1.36$	$t = 0.001$	$C = 0.09$
nylon	$\alpha = 0.26$	$\rho = 0.92$	$t = 1.46$	$C = 0.01$

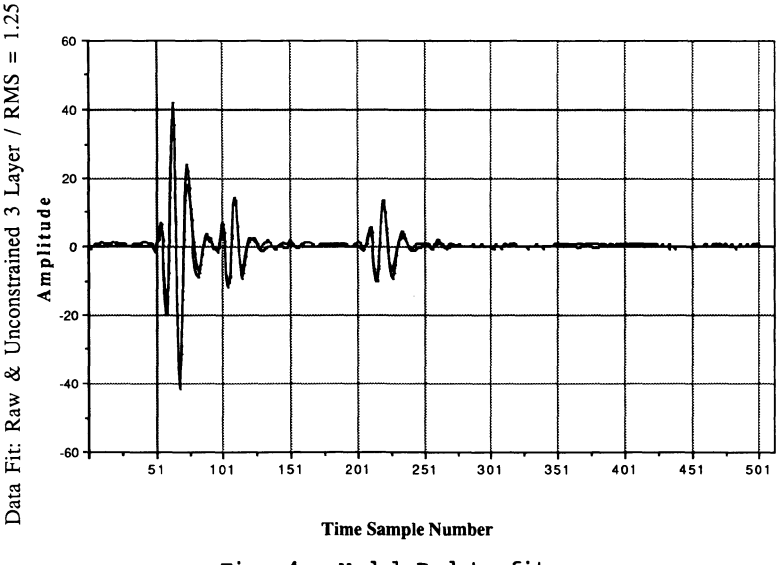

Fig. 4. Model B data fit.

Model C: A two layer model with no provision for the inclusion of a water layer (bondline perhaps). The measured values were used for the initial guess parameters. No constraints, aside from nonnegativity constraints (i.e., parameters cannot be negative) and the natural constraints imposed by our particular choice of forward solution, were applied. Convergence and data fit are shown in Figs. 2 and 5. The data fit is virtually the same as the three layer models. This fit was obtained by the algorithm adjustment of the densities of the two layers to obtain the proper impedance contrast at the Lucite / nylon interface. The final model was:

Lucite	$\alpha = 0.27$	$\rho = 1.04$	$t = 0.56$	$C = 0.000$
nylon	$\alpha = 0.26$	$\rho = 0.85$	$t = 1.46$	$C = 0.006$

Model D: A two layer model similar to Model C except that a very poor initial guess was used. Convergence and data fit are shown in Figs. 2 and 6. The final model was:

Lucite	$\alpha = 0.41$	$\rho = 0.69$	$t = 0.85$	$C = 0.000$
nylon	$\alpha = 0.26$	$\rho = 0.79$	$t = 1.45$	$C = 0.000$

CONCLUSIONS

Material property estimates will, in general, be nonunique. We must find some way to, first, determine that nonuniqueness, and hence, develop techniques to reduce it. Material property estimation is the result of a data acquisition procedure. As such, it is important to realize the statistical nature of the exercise.

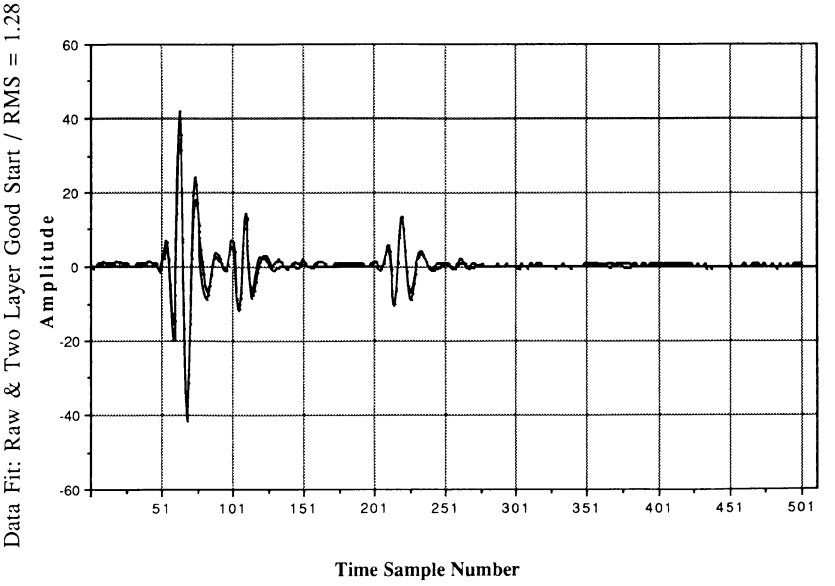

Fig. 5. Model C data fit.

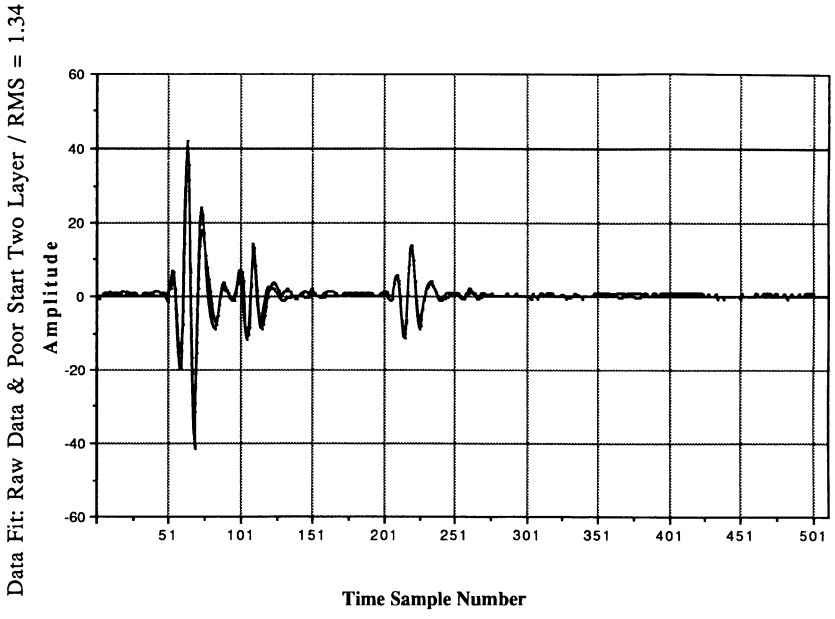

Fig 6. Model D data fit.

We have shown four different models which were all capable of fitting a pulse echo A-Scan. This illustrates the inherent nonuniqueness in this type of data. Through transmission data would display similar characteristics. Nonlinear inversion was shown to be capable of determining best fit model parameters for each model and was able, via the parameter covariances and correlation coefficients, to provide confidence intervals within a linear range of the best fit model parameters. These confidence intervals provide an estimate of uniqueness.

Two steps can be taken to reduce the nonuniqueness of, in this example, pulse echo data. One would be to have very good control on the material properties of the object being tested and simply constrain them. This presents obvious drawbacks. A better approach would be to use more sophisticated transmitter / receiver geometries.

REFERENCES

1. Y. Bard, Nonlinear Parameter Estimation (Academic Press, Inc., 1974).
2. W. Menke, Geophysical Data Analysis: Discrete Inverse Theory (Academic Press, Inc., 1984).
3. D. W. Marquardt, An Algorithm for Least Squares Estimation on Nonlinear Parameters, J. Soc. Ind. Appl. Math., 1963, No. 2, pp. 431-441.